Use of Selenium Concentration in Whole Blood, Serum, Toenails, or Urine as a Surrogate Measure of Selenium Intake

Matthew P. Longnecker,[1] Daniel O. Stram,[2] Philip R. Taylor,[3] Orville A. Levander,[4] Marmion Howe,[5] Claude Veillon,[4] Patricia A. McAdam,[4] Kristine Y. Patterson,[4] Joanne M. Holden,[6] J. Steven Morris,[7] Christine A. Swanson,[8] and Walter C. Willett[9]

We examined the validity of using the selenium level in a single biological specimen as a surrogate measure of usual intake. We used data from 77 free-living adults from South Dakota and Wyoming. Subjects provided multiple 1-day duplicate-plate food composites, repeated specimens of blood and toenails, and 24-hour urine collections. We developed a statistical calibration method that incorporated measurement error correction to analyze the data. The Pearson correlation coefficients between selenium intake and a single selenium status measure, after deattenuation to adjust for the effect of within-person variation in intake, were: 0.78 for whole blood, 0.74 for serum, 0.67 for toenails, and 0.86 for urine. We present formulas to estimate the intake of individuals, based on selenium levels in a single specimen of blood, toenails, or urine. In these data, the concentration of selenium in a single specimen of whole blood, serum, or toenails served reasonably well as a measure for ranking subjects according to long-term selenium intake but provided only a rough estimate of intake for each subject. (Epidemiology 1996;7:384–390)

Keywords: human, measurement error, nutrition assessment, selenium, calibration, exposure assessment.

Selenium is a trace element with well defined deficiency and toxicity states.[1] Selenium is a component of glutathione peroxidase, an enzyme with antioxidant activity, and the relation of selenium status to risk of heart disease and cancer is under investigation.[2,3]

The estimation of intake of selenium from self-reported food intake and from the average selenium content of foods, which is the usual method of assessing nutrient intake, is not accurate because the selenium content of any particular food item varies with the selenium concentration in the soil where the food was produced.[4] Earlier studies have established that the concentration of selenium in toenails and blood and the amount in urine reflect the levels of dietary selenium intake.[5–12] None of these studies, however, has measured selenium intake directly and related this intake to the level of selenium in a single specimen of blood, nails, or urine. In epidemiologic studies, the level of selenium in only one biological specimen, such as a set of toenails, has typically been used as an exposure measure.[2,3]

In this report, we examine the validity of using the selenium concentration in a single biological specimen as a surrogate measure of selenium intake. In addition, we present formulas that, along with a single measure of selenium in a biological specimen and other information that may be available, can be used to estimate the selenium intake of an individual. Although we had multiple observations for diet and other measures for each subject, we used a statistical procedure to estimate the relation between a single measure of selenium in blood, toenails, or urine and true long-term intake. We corrected correlation and regression coefficients for measurement error due to within-person variation in intake. The statistical methods that we developed may be of interest to those concerned with calibration methods for measurement error correction in exposure assessment.

From the [1]Department of Epidemiology, UCLA School of Public Health, Los Angeles, CA; [2]Division of Biostatistics, Department of Preventive Medicine, USC School of Medicine, Los Angeles, CA; [3]Cancer Prevention Studies Branch, Division of Cancer Prevention and Control, National Cancer Institute, Rockville, MD; [4]Nutrient Requirements and Functions Laboratory and [6]Nutrient Composition Laboratory, U.S. Department of Agriculture, Beltsville, MD; [5]South Dakota School of Mines and Technology, Rapid City, SD; [7]Research Reactor Facility, University of Missouri, Columbia, MO; [8]Environmental Epidemiology Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD; and [9]Departments of Nutrition and Epidemiology, Harvard School of Public Health; Channing Laboratory, Harvard Medical School; and Department of Medicine, Brigham and Womens' Hospital, Boston, MA.

Address reprint requests to: Matthew P. Longnecker, NIEHS Epidemiology Branch, P.O. Box 12233, MD A3-05, Research Triangle Park, NC 27709.

Matthew P. Longnecker was the recipient of a Junior Faculty Research Award from the American Cancer Society when this manuscript was prepared.

Submitted October 2, 1995; final version accepted March 26, 1996.

Subjects and Methods

We used data from a study of the safety of high dietary intake of selenium.[12] Seventy-seven subjects from western South Dakota and eastern Wyoming were selected either from telephone books (N = 30, selected at random) or from areas where unusually high selenium intakes were suspected (N = 47; see Longnecker *et al*[12] for details of the selection procedure). One adult per household was included in the present study.

Subjects were free-living, and each was studied over either a 6- or 12-month period. Once per season, for 2 days, each subject saved duplicate portions of all foods and beverages consumed. One quarter of the food collections for each subject were on Saturday or Sunday. Twenty-four-hour urine samples were obtained on the second day of food collection for the season. Blood specimens were drawn the morning after the second day of food collection, following an overnight fast. Subjects' toenails were collected at the time of each blood drawing. By design, for 45 subjects, data for four seasons were collected, and for the remaining 32 subjects, data for just two seasons were collected (summer and winter). Thus, 45 subjects were instructed to collect food for 8 days and to provide four blood, nail, and urine specimens; 32 subjects were asked to collect food for 4 days and to provide two blood, nail, and urine specimens. Self-administered questionnaires were used to ascertain age, gender, height, weight, and smoking status. Subjects also completed a detailed semiquantitative food frequency questionnaire, which inquired about the frequency of intake of 116 food items and the use of dietary supplements. The validity of this instrument in this population was reported earlier.[13] The study was conducted in 1985–1987.

The duplicate food portions were kept refrigerated or frozen until homogenized at Rapid City Regional Hospital in Rapid City, SD. Aliquots of the food homogenates were frozen at −9°C and sent via overnight mail on dry ice to the Human Nutrition Research Center of the U.S. Department of Agriculture in Beltsville, MD. Fasting venous blood samples were drawn into trace element-free Vacutainers (Becton Dickinson, Lincoln, NJ). Aliquots of 24-hour urine collections, whole blood, and serum were frozen in Rapid City and sent to Beltsville. Nail samples (clippings from all toes) from each subject were placed in manila envelopes and mailed to Boston, where they underwent ultrasonic cleaning in a bath of distilled water, were dried, and then were sent to the University of Missouri at Columbia.

The selenium content of serum, whole blood, and urine samples and of the subjects' diet was determined using an isotope dilution technique and gas chromatographic-mass spectrometric analysis.[14] The analytical blank for all samples averaged 3.8 pmol Se. The detection limit of the method [defined as the mean of the analytical blank plus three times the standard deviation (SD)] was 6.4 pmol Se. Quality control was maintained by regular determinations of in-house reference pools of serum, whole blood, and urine. The coefficient of variation for duplicate analyses of the in-house reference pools was always less than 2%.

The selenium content of toenail specimens was determined by neutron activation analysis.[15] The concentration of selenium in the National Institute of Standards and Technology Standard Reference Material #1577 (bovine liver), determined by neutron activation analysis, was 14.1 ± 0.6 μmol per kg (mean ± SD), as compared with a certified value of 14 ± 1 μmol per kg. The coefficient of variation for 38 repeated measurements was 4.5%.

Statistical Methods

Our interest was in the relation between a single measure of selenium status and the subjects' true long-term selenium intake. Our dataset, however, had multiple observations of selenium levels in blood, toenails, and urine. To make use of all of these observations while obtaining results for one measure, we used the linear regression methods described in Appendix A. These methods also allowed us to deattenuate correlation coefficients for error due to day-to-day variation (within-person error) in selenium intake and within-subject variation in measures of selenium status, and to deattenuate regression coefficients for variation (within-person error) in measures of selenium status.[16]

The distributions of the measures of selenium were skewed, with the long tail to the right. We therefore used a $\log_e$ transformation for these variables before performing the correlation and regression analyses.

In addition to fitting univariate models of selenium intake as a function of selenium level in blood, toenails, or urine, we also fit multivariate models to incorporate information from other variables that improved prediction of intake. We identified these covariates in standard linear regression models, using the means of subjects' values of intake and of their blood, nails, or urine. We identified specific covariates by first considering whether any among a candidate list of nondietary covariates (age, gender, smoking, and estimated lean body mass[17]) improved the prediction of selenium intake. We started with models that included a measure of selenium status and all of these variables and then deleted the variables that did not improve prediction; the result was a set of parsimonious models for prediction of intake in the absence of dietary data. We then considered models of selenium intake with one of the selenium indices as a predictor variable, as well as the dietary variables (energy, methionine, and alcohol), and the other candidate covariates listed above, and we again identified parsimonious models. We selected these dietary variables because of the correlation of energy intake with intake of many micronutrients,[16] the possibility that dietary methionine might predict intake of selenium (as selenomethionine), and reports that alcohol intake may affect selenium metabolism.[18] The covariates were incorporated into the models described in Appendix A in the iteratively reweighted least-squares step.

TABLE 1. Mean Daily Selenium Intake, Urinary Excretion, and Selenium Concentration in Serum, Whole Blood, and Toenails*

	Mean ± SD (N = 77)	Range
Intake (μmol/day)	2.99 ± 1.81	0.79–9.16
Whole blood (μmol/kg)	3.94 ± 1.33	2.37–7.99
Serum (μmol/liter)	2.45 ± 0.68	1.56–4.58
Toenails (μmol/kg)	19.2 ± 7.0	10.5–47.2
Urine (μmol/day)	2.05 ± 1.38	0.32–6.09

* One micromole of selenium weighs 79.0 μg. We calculated these means by first determining the within-person mean of all replicate values and then averaging these across subjects, with all subjects equally weighted.

Results

The mean age of the 77 subjects was 50 years (SD = 14 years; range = 22–81 years). Forty-one subjects were men, and 36 were women. Eighteen subjects (23%) were current smokers. The mean weight of the subjects was 73.2 kg (SD = 13.2 kg; range = 46.8–104.5 kg). The mean daily energy intake was 9,800 kJ. All subjects were white, and none took selenium supplements. The average selenium intake in this population (2.99 μmol per day) exceeded the estimated average daily selenium intake in the United States, 0.8–2.0 μmol[19–21] (Table 1).

The intraclass correlation coefficient among repeated measures of selenium intake and measures of selenium status was high for blood and toenails and moderately high for urine and intake (Table 2). We examined the simple correlation between the mean of each subject's selenium intake and the mean level of selenium in the biological measures and found that, when selenium intake was expressed per kg body weight, these correlations were greater for all measures except urine (Table 2). Therefore, in subsequent analyses of selenium levels in whole blood, serum, or toenails, we expressed selenium intake per kg body weight.

When we examined the correlation between selenium intake adjusted for measurement error and a single measure of selenium status, the correlations ranged from 0.67 to 0.86, indicating that a single measure of selenium status served reasonably well for ranking subjects according to long-term intake (Table 3). With adjustment for gender, or gender and age, the correlations in Table 3 increased slightly.

TABLE 2. Intraclass Correlation Coefficients among Replicate Measures and Pearson Correlation Coefficients between Mean log Selenium Measure and Mean log Selenium Intake (or Mean log Selenium Intake/kg) (N = 77)

Measure of Se Intake or Status	Intraclass Correlation	Pearson Correlations Mean Intake (μmol/day) log(Se)	log(Se)/kg
log(Se intake)/kg (μmol/day)	0.62	0.96	1.00
log(Se intake) (μmol/day)	0.65	1.00	0.96
log(whole blood) (μmol/kg)	0.95	0.70	0.75
log(serum) (μmol/liter)	0.88	0.64	0.73
log(toenails) (μmol/kg)	0.91	0.62	0.65
log(urine) (μmol/day)	0.77	0.89	0.84

TABLE 3. Correlation Coefficients between Adjusted Selenium Intake/kg and a Single Measure of Selenium (Se) Status in 77 Subjects*

Measure of Se Status	log(Se Intake)/kg Adjusted for: Error in Measurement of Intake	Error in Measurement of Intake and Specified Factors†
log(whole blood) (μmol/kg)	0.78	0.81 Gender
log(serum) (μmol/liter)	0.74	0.79 Gender
log(toenails) (μmol/kg)	0.67	0.71 Gender, age
log(urine) (μmol/day)‡	0.86	0.87 Gender

* Selenium intake is in μmol/day.
† The covariates included in the partial correlation analysis appear in the column below. The covariates are those identified as important in the multivariate models that considered nondietary variables.
‡ Prediction of selenium intake with urine values was of log(Se) intake per day, not per kg.

Selenium intake correlated strongly with energy intake estimated from the food frequency questionnaire (Table 4). Gender was less strongly associated with selenium intake, and when intake was expressed per kg body weight, the association decreased. Age was inversely related to energy intake. Correlations for other variables were not important for prediction of selenium intake; these results are not shown.

Given one measure of selenium status, the univariate regression coefficients that may be used to estimate selenium intake for a subject are shown in Table 5. Table 5 also shows the coefficients for prediction given one measure of selenium status and selected nondietary information, and it shows the coefficients for prediction given one measure of selenium status and selected dietary and nondietary information. The R^2 values for the models shown in Table 5 were generally in the range of 0.50–0.75.

We demonstrate use of these equations as follows: suppose a 70-kg man consuming 8,000 kJ per day was found to have a serum selenium value of 2.0 μmol per liter. The estimated selenium intake, using the appropriate formula from Table 5, would be exp[70 kg × [$-1.31 \times 10^{-2} + 2.00 \times 10^{-2} \times$ log(2.0 μmol per liter) + $8.27 \times 10^{-7} \times$ 8,000 kJ]], or 1.68 μmol per day [95% confidence interval (CI) = 1.00–2.81]. We present the data necessary for calculating prediction standard errors

TABLE 4. Pearson Correlation Coefficients between Mean log Selenium Intake and Selected Characteristics in 77 Subjects*

	Energy Intake (kJ)	Gender	Age (Years)
log(Se intake)	0.59	−0.38	−0.03
log(Se intake)/kg	0.53	−0.23	−0.08
Energy intake		−0.27	−0.17
Gender			−0.17

* Units for selenium intake are μmol/day. Gender = 0 if male, 1 if female. Spearman correlations among these variables were nearly the same as the values shown.

TABLE 5. Regression Coefficients [and standard errors (SE)] for Models Predicting the log of Selenium Intake/kg [log (μmol × day^{-1}) × kg^{-1}]; the Models Were Fitted to Reflect the Relation of Long-Term Intake to One Observation of the Independent Variable

Independent Variable	Coefficient	(SE)	R^2
Univariate models			
Intercept	-1.39×10^{-2}	(2.72×10^{-3})	0.61
log(whole blood) (μmol/kg)	1.93×10^{-2}	(2.03×10^{-3})	
Intercept	-7.03×10^{-3}	(2.30×10^{-3})	0.55
log(serum) (μmol/liter)	2.16×10^{-2}	(2.63×10^{-3})	
Intercept	-3.33×10^{-2}	(6.51×10^{-3})	0.45
log(toenails) (μmol/kg)	1.55×10^{-2}	(2.25×10^{-3})	
Intercept	5.32×10^{-1}	(3.62×10^{-2})	0.74
log(urine) (μmol/day)*	6.70×10^{-1}	(5.15×10^{-2})	
Multivariate models that include nondietary variables			
Intercept	-1.21×10^{-2}	(2.55×10^{-3})	0.66
log(whole blood) (μmol/kg)	1.93×10^{-2}	(1.88×10^{-3})	
Gender	-3.98×10^{-3}	(1.10×10^{-3})	
Intercept	-5.39×10^{-3}	(2.08×10^{-3})	0.63
log(serum) (μmol/liter)	2.23×10^{-2}	(2.35×10^{-3})	
Gender	-4.96×10^{-3}	(1.11×10^{-3})	
Intercept	-2.73×10^{-2}	(6.31×10^{-3})	0.51
log(toenails) (μmol/kg)	1.59×10^{-2}	(2.08×10^{-3})	
Age (years)	-8.24×10^{-5}	(4.64×10^{-5})	
Gender	-4.02×10^{-3}	(1.29×10^{-3})	
Intercept	6.04×10^{-1}	(4.86×10^{-2})	0.75
log(urine) (μmol/day)*	6.40×10^{-1}	(5.18×10^{-2})	
Gender	-1.20×10^{-1}	(6.05×10^{-2})	
Multivariate models that include energy intake and nondietary variables			
Intercept	-1.92×10^{-2}	(2.66×10^{-3})	0.69
log(whole blood) (μmol/kg)	1.73×10^{-2}	(1.71×10^{-3})	
Gender	-2.61×10^{-3}	(1.01×10^{-3})	
Energy intake (kJ)	9.26×10^{-7}	(1.91×10^{-7})	
Intercept	-1.31×10^{-2}	(2.40×10^{-3})	0.66
log(serum) (μmol/liter)	2.00×10^{-2}	(2.15×10^{-3})	
Gender	-3.61×10^{-3}	(1.02×10^{-3})	
Energy intake (kJ)	8.27×10^{-7}	(1.92×10^{-7})	
Intercept	-3.56×10^{-2}	(5.84×10^{-3})	0.52
log(toenails) (μmol/kg)	1.40×10^{-2}	(1.90×10^{-3})	
Age (years)	-3.42×10^{-5}	(4.25×10^{-5})	
Gender	-2.31×10^{-3}	(1.21×10^{-3})	
Energy intake (kJ)	1.03×10^{-6}	(2.25×10^{-7})	
Intercept	3.56×10^{-1}	(1.21×10^{-1})	0.76
log(urine) (μmol/day)*	5.93×10^{-1}	(5.76×10^{-2})	
Gender	-1.20×10^{-1}	(5.95×10^{-2})	
Energy intake (kJ)	6.76×10^{-7}	(1.22×10^{-5})	

* Dependent variable is log(Se intake), not per kg. Gender is coded as 1 for female, 0 otherwise.

in Appendix B and for calculating the confidence intervals for predicted intakes in Appendix C.

Finally, we evaluated whether using replicate measures of selenium status provided an improved estimate of selenium intake compared with that obtained using a single measure of status. We found that use of replicate measures added little (results not shown). For example, the increase in error-adjusted correlation coefficients (between intake and surrogate measure) due to increasing the number of observations to four from one were: whole blood, 0.02; serum, 0.04; toenails, 0.02; and urine, 0.08.

Discussion

The correlation coefficients corrected for error in intake indicate that use of one surrogate measure to rank subjects according to long-term intake will yield reasonably accurate results. Whole blood and serum were about equally good measures of long-term intake, and toenails did nearly as well. Relative to the validity of other accepted measures of dietary intake,[13] the selenium level in toenails, serum, and whole blood as surrogate measures of intake work well. Because the intraclass correlation coefficients for repeated measures of selenium status were all fairly high in our study, the error-adjusted correlations between mean individual intakes and one surrogate measure of selenium intake (Table 3) were similar to those found when the average of several surrogate measures was used (Table 2).

The regression coefficients presented, used in conjunction with a single measure of selenium status, will provide only a rough estimate of the intake of an individual, probably because of appreciable individual variation in response to a given intake.[11] Consideration of energy intake or other information in addition to a selenium status measure resulted in a slightly more precise estimate of intake, although the 95% confidence intervals were still wide. A larger study size would not have resulted in equations giving substantially more precise estimates of an individual's intake, because the dominant source of error in predicted selenium intake for a single individual is the between-person variance in true long-term average of the surrogate measure, which is not affected by study size (Appendix C, last part).

In these data, the correlation between selenium intake and the amount of selenium in one 24-hour urine specimen was greater than for other biological measures. The correlation was especially high because one-half of the urine specimens were collected on the same day as duplicate-plate food specimens. The rate of excretion of a given dose of selenium is greatest within 24 hours of ingestion.[22] The ratio of within-to-between person variation in urine selenium (0.30) was higher than for other surrogate measures (serum, 0.14; whole blood 0.05; toenails, 0.10). Thus, a single specimen of one of the other measures is likely to be a better reflection of long-term intake.

Energy intake is correlated with the intake of many micronutrients.[16] Thus, it was not surprising that energy intake should be predictive of selenium intake. In an earlier, related analysis of a subset of these data, Swanson *et al*[5] found that selenium intakes among women (per kg body weight) were 23% lower than among men ($P = 0.05$); our data confirm a lower intake among women. Tissue levels of selenium, however, were not much different between the sexes in Swanson *et al*'s analysis. The higher selenium intake of men, with tissue selenium levels similar to that of women, was thought to reflect a greater selenium need among men, for whom the pro-

portion of lean body mass is greater. In the present analysis, however, when estimated lean body mass and gender were both considered as potential predictor variables, only gender was important. Gender may better reflect true lean body mass than our estimate, based on age, gender, height, and weight. Or gender may carry data about true energy intake in addition to that estimated by the food frequency questionnaire. Another possibility is that some as yet unrecognized difference between the sexes in selenium metabolism explains the discrepancy between selenium intake and tissue levels of men and women. The percentage of selenium excreted in urine, however, was similar in men and women (69% *vs* 66%). Prediction models (like those shown in Table 5) that also included a term for interaction between selenium tissue levels and gender fit the data better only for toenails; the results revealed that the gender difference in intake (males higher) predicted for a given level of toenail selenium was more pronounced at lower nail selenium levels and negligible at higher nail levels.

The equations developed might be less applicable in another population owing to variation in selenium metabolism,[11] or where the chemical forms of selenium in the diet differed from those in the South Dakota and Wyoming subjects. The average daily selenium excretion in urine was 69% of average intake in this population (Table 1). Levander *et al*[23] and Yang *et al*[6] noted that, in their subjects, the proportion excreted was 40–60%. Perhaps a greater proportion of dietary selenium in South Dakota and Wyoming was inorganic and, thus, less well retained in the body.[24] Analysis of duplicate portions gives the most valid estimate of nutrient intake available.[25] The method probably causes underestimates of intake[26] and may cause misclassification.[25] The error thus incurred may cause our estimates of selenium intake to be slightly low and causes an unquantifiable but likely small amount of imprecision.

Acknowledgments

We are grateful to Kathleen Corr for her assistance in enrolling subjects and collecting data and to Ed Kushman for technical assistance. Oscar E. Olson guided us to several of the ranchers who participated in the study.

References

1. Levander OA. A global view of human selenium nutrition. Annu Rev Nutr 1987;7:227–250.
2. Kok FJ, Hofman A, Witteman JC, de Bruijn AM, Kruyssen DH, de Bruijn M, Valkenburg HA. Decreased selenium levels in acute myocardial infarction. JAMA 1989;261:1161–1164.
3. van den Brandt PA, Goldbohm RA, van't Veer P, Bode P, Dorant E, Hermus RJ, Sturmans F. Toenail selenium levels and the risk of breast cancer. Am J Epidemiol 1994;140:20–26.
4. Levander OA. The need for measures of selenium status. J Am Coll Toxicol 1986;5:37–44.
5. Swanson CA, Longnecker MP, Veillon C, Howe M, Levander OA, Taylor PR, McAdam PA, Brown CC, Stampfer MJ, Willett WC. Relation of selenium intake, age, gender, and smoking to indices of selenium status of adults residing in a seleniferous area. Am J Clin Nutr 1990;52:858–862.
6. Yang G, Zhou R, Yin S, Gu L, Yan B, Liu Y, Liu Y, Li X. Studies of safe maximal daily dietary selenium intake in a seleniferous area in China. I. Selenium intake and tissue selenium levels of the inhabitants. J Trace Elem Electrolytes Health Dis 1989;3:77–87.
7. Schrauzer GN, White DA. Selenium in human nutrition: dietary intakes and effects of supplementation. Bioinorg Chem 1978;8:303–318.
8. Hunter DJ, Morris JS, Chute CG, Kushner E, Colditz GA, Stampfer MJ, Speizer FE, Willett WC. Predictors of selenium concentration in human toenails. Am J Epidemiol 1990;132:114–122.
9. van den Brandt PA, Goldbohm RA, van't Veer P, Bode P, Hermus RJ, Sturmans F. Predictors of toenails selenium levels in men and women. Cancer Epidemiol Biomarkers Prev 1993;2:107–112.
10. Ovaskainen ML, Virtamo J, Alfthan G, Haukka J, Pietinen P, Taylor PR, Huttunen JK. Toenail selenium as an indicator of selenium intake among middle-aged men in an area with low soil selenium. Am J Clin Nutr 1993;57:662–665.
11. Longnecker MP, Stampfer MJ, Morris JS, Spate V, Baskett C, Mason M, Willett WC. A 1-y trial of the effect of high-selenium bread on selenium levels in blood and toenails. Am J Clin Nutr 1993;57:408–413.
12. Longnecker MP, Taylor PR, Levander OA, Howe M, Veillon C, McAdam PA, Patterson KY, Holden JM, Stampfer MJ, Morris JS, Willett WC. Selenium in diet, blood, and toenails in relation to human health in a seleniferous area. Am J Clin Nutr 1991;53:1288–1294.
13. Longnecker MP, Lissner L, Holden JM, Flack VF, Taylor PR, Stampfer MJ, Willett WC. The reproducibility and validity of a self-administered semiquantitative food frequency questionnaire in subjects from South Dakota and Wyoming. Epidemiology 1993;4:356–365.
14. Reamer DC, Veillon C. A double isotope dilution method for using stable selenium isotopes in metabolic tracer studies: analysis by gas chromatography/mass spectrometry (GC/MS). J Nutr 1983;113:786–792.
15. Morris JS, Stampfer MJ, Willett W. Dietary selenium in humans: toenails as an indicator. Biol Trace Elem Res 1983;5:529–537.
16. Willett WC. Nutritional Epidemiology. New York: Oxford University Press, 1990.
17. Watson PE, Watson ID, Batt RD. Total body water volumes for adult males and females estimated from simple anthropometric measurements. Am J Clin Nutr 1980;33:27–39.
18. Lloyd B, Lloyd RS, Clayton BE. Effects of smoking, alcohol, and other factors on the selenium status of a healthy population. J Epidemiol Community Health 1983;37:213–217.
19. Mahaffey KR, Corneliussen PE, Jelinek CF, Fiorino JA. Heavy metal exposure from foods. Environ Health Perspect 1975;12:63–69.
20. Pennington JAT, Young BE, Wilson DB, Johnson RD, Vanderveen JE. Mineral content of foods and total diets: The Selected Minerals in Foods Survey, 1982 to 1984. J Am Diet Assoc 1986;86:876–891.
21. Pennington JA, Young BE, Wilson DB, Johnson RD, Vanderveen JE. Nutritional elements in U.S. diets: results from the total diet study, 1982–1986. J Am Diet Assoc 1989;89:659–664.
22. Swanson CA, Patterson BH, Levander OA, Veillon C, Taylor PR, Helzlsouer K, McAdam PA, Zech LA. Human [^{74}Se]selenomethionine metabolism: a kinetic model. Am J Clin Nutr 1991;54:917–926.
23. Levander OA, Sutherland B, Morris VC, King JC. Selenium balance in young men during selenium depletion and repletion. Am J Clin Nutr 1981;34:2662–2669.
24. Alfthan G, Aro A, Arvilommi H, Huttunen JK. Selenium metabolism and platelet glutathione peroxidase activity in healthy Finnish men: effects of selenium yeast, selenite, and selenate. Am J Clin Nutr 1991;53:120–125.
25. West CE, van Staveren WA. Food composition, nutrient intake, and the use of food composition tables. In: Margetts BM, Nelson M, eds. Design Concepts in Nutritional Epidemiology. New York: Oxford University Press, 1991;102–103.
26. Basiotis PP, Thomas RG, Kelsay JL, Mertz W. Sources of variation in energy intake by men and women as determined from one year's daily dietary records. Am J Clin Nutr 1989;50:448–453.
27. Rosner BA, Willett WC, Speigelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 1989;8:1051–1069.

Appendix A

Detailed Description of the Statistical Methods

The goal of the statistical methods is to predict the value of the long-term mean of the log of daily selenium per kg for subject i, $\bar{y}_i^{\infty}$, based upon one or more measurements of each of the surrogate measurements, z_{ij}. We estimated the parameters in a linear prediction equation for estimating $\bar{y}_i^{\infty}$ using its expected value $E(\bar{y}_i^{\infty}|\bar{z}_i^{k})$, conditional upon having observed k z_{ij}, for each subject i. Denoting the mean of these observations, $\Sigma_{j=1}^{k} z_{ij}$, as $\bar{z}_i^{k}$, we estimate the intercepts, a^k, and slopes, b^k, in the equation:

$$E(\bar{y}_i^{\infty}|\bar{z}_i^{k}) = a^k + b^k \bar{z}_i^{k}. \quad (1)$$

The a^k and b^k depend upon the number of observations k, because the amount of random measurement error in the estimate, $\bar{z}_{i.}^k$, of a subject's surrogate measure depends upon k. As described by Rosner *et al*,[27] such measurement error has the effect of biasing estimates of slopes and intercepts, so that estimates which are obtained using one value of k would not, in general, be applicable in predicting $\bar{y}_i^\infty$ when some other number of surrogate measurements are utilized. We first estimate the parameters in the linear regression:

$$E(\bar{y}_i^\infty|\bar{z}_{i.}^\infty) = a^\infty + b^\infty \bar{z}_{i.}^\infty \quad (2)$$

as well as the residual variance, $Var(\bar{y}_i^\infty \mid \bar{z}_{i.}^\infty) \equiv \sigma^2$, and then reconstruct a^k and b^k according to the formulas:

$$a^k = a^\infty + b^\infty(1 - R^k) \quad (3)$$

and

$$b^k = b^\infty R^k \quad (4)$$

where

$$R^k = \frac{v^2}{v^2 + \sigma_\epsilon^2/k}. \quad (5)$$

Here, v^2 is the between-person variance of the true long-term average of the surrogate value, $\bar{z}_{i.}^\infty$, and σ_ϵ^2 is the estimate of the day-to-day variation (within-person) in an individual's surrogate measurements. The variances v^2 and σ_ϵ^2 were estimated by fitting a mixed-effects model for the surrogate values z_{ij}, which also provided an estimate of the population mean of the $\bar{z}_{i.}^\infty$, which is denoted z^∞.

Since the $\bar{z}_{i.}^\infty$ in Eq 2 were unknown, they were approximated as:

$$\hat{\bar{z}}_{i.}^\infty \equiv E(\bar{z}_{i.}^\infty|\bar{z}_{i.}^{m_i}) = \bar{z}^\infty + R^{m_i}(\bar{z}_{i.}^{m_i} - \bar{z}^\infty),$$

where m_i is the number of surrogate measurements available for subject i. Finally, these $\hat{\bar{z}}_{i.}^\infty$ were used to estimate a^∞ and b^∞ and σ^2, in an iteratively reweighted least-squares regression (IRWLS) of

$$\bar{y}_i^{n_i} = 1/n_i \sum_{i=1}^{n_i} y_i$$

on $\hat{\bar{z}}_{i.}^\infty$, where n_i is the number of days of dietary selenium measurements for subject i, and the weights used in the regression corresponded to the inverse of

$$Var(\bar{y}_{i.}^{n_i}|\hat{\bar{z}}_{i.}^\infty) = (b^\infty)^2(1 - R^{m_i})\sigma_\epsilon^2 + \sigma_\epsilon^2/n_i + \sigma^2.$$

In the IRWLS regression, initial estimates of the unknown parameters σ^2 and b^∞ were used to calculate initial weights, and these were subsequently updated in each iteration of the regression. The other parameters in

$$Var(\bar{y}_{i.}^{n_i}|\hat{\bar{z}}_{i.}^\infty)$$

were replaced with their estimates from the fit of mixed-effects models for both y_{ij} and z_{ij}, with σ_ϵ^2 representing the day-to-day variation in intake, y_{ij}. Finally, a^k and b^k were estimated from Eq 3 and Eq 4.

Appendix B

Estimates Needed to Calculate the Variance of Predicted Selenium Intakes

TABLE A1. Parameter Estimates from Mixed-Effects Models ($y_{ij} = \bar{y}_i^\infty + e_{ij}$) for Dependent Variables*

	$\bar{y}^\infty$	v^2	σ_e^2
log(dietary selenium)/kg	0.0114	5.79×10^{-5}	3.50×10^{-5}
log(dietary selenium)	0.452	0.398	0.118

* The first column of values is population means, the second is between-individual variances, and the third is within-individual variances.

TABLE A2. Parameter Estimates from Mixed-Effects Models ($z_{ij} = \bar{z}_{i.}^\infty + \epsilon_{ij}$) for Independent Variables*

	$\bar{z}^\infty$	v^2	σ_ϵ^2
log(serum selenium)	0.857	0.0599	0.00824
log(whole blood selenium)	1.32	0.0902	0.00500
log(toenail selenium)	2.90	0.0993	0.00959
log(urine selenium)	0.452	0.398	0.118

* The first column of values is population means, the second is between-individual variances, and the third is within-individual variances.

TABLE A3. Iteratively Reweighted Least-Squares Estimates of Regression Parameters a^∞, b^∞, and σ^2 in the Equations $E(\bar{y}_i^\infty|\bar{z}_{i.}^\infty) = a^\infty + b^\infty\bar{z}_{i.}^\infty$, $Var(\bar{y}_i^\infty|\bar{z}_{i.}^\infty) = \sigma^2$

	a^∞	b^∞	σ^2
Basic model			
Serum	−0.00958	0.0246	2.17×10^{-5}
Whole blood	−0.0153	0.0204	2.05×10^{-5}
Toenail	−0.0376	0.0170	$2.9? \times 10^{-5}$
Urine	0.442	0.868	1.39×10^{-2}
Model including nondietary covariates			
Serum	−0.00801	0.0253	1.56×10^{-5}
Whole blood	−0.0135	0.0204	1.64×10^{-5}
Toenail	−0.0318	0.0174	2.42×10^{-5}
Urine	0.519	0.829	7.54×10^{-3}
Model with all covariates			
Serum	−0.0155	0.0227	1.04×10^{-5}
Whole blood	−0.0205	0.0182	1.11×10^{-5}
Toenail	−0.0396	0.0154	1.75×10^{-5}
Urine	0.276	0.768	4.63×10^{-3}

TABLE A4. Covariances for $\hat{a}^1$, $\hat{b}^1$

log(whole blood selenium)/kg	-5.38×10^{-6}
log(serum selenium)/kg	-5.83×10^{-6}
log(toenail selenium)/kg	-1.45×10^{-5}
log(urine)/day	-1.04×10^{-3}

Appendix C

Prediction Errors for Estimating Long-Term Dietary Selenium Intake

The residual variance in true long-term (transformed) dietary selenium, $\bar{y}_i^\infty$, given the average of k measurements of the surrogate values, z_{ij}, may be written as:

$$Var(\bar{y}_i^\infty|\bar{z}_{i.}^k) = \sigma^2 + b^\infty(b^\infty - b^k)v^2. \quad (6)$$

This variance may be estimated from the values given in the tables in Appendix B, for each of the different surrogate values. An approximate 95% confidence region for this prediction is based on:

$$\bar{y}_i^{\infty} \pm 1.96\sqrt{Var(\bar{y}_i^{\infty}|\bar{z}_i^1)}.$$

For example, suppose a 70-kg subject was found to have a serum selenium value of 2.0 μmol per liter, and hence a predicted selenium intake of exp{70 kg × [−7.03 × 10^{-3} + 2.16 × 10^{-2} × log(2 μmol per liter)]} = 1.74 μmol per day. Because for serum selenium we have $\hat{\sigma}^2 = 2.17 \times 10^{-5}$, $\hat{b}^{\infty} = 2.46 \times 10^{-2}$, $\hat{b}^1 = 2.16 \times 10^{-2}$, and $\hat{\sigma}^2 = 0.0599$, the estimate of

$$\bar{y}_i^{\infty} \pm 1.96\sqrt{Var(\bar{y}_i^{\infty}|\bar{z}_i^1)}$$

is 5.11 × 10^{-3}. Thus, the 95% CI for dietary selenium intake is exp{70 kg × [−7.03 × 10^{-3} + 2.16 × 10^{-2} × log(2 μmol per liter) ± 1.96 × 5.11 × 10^{-3}]} = 0.862–3.51 μmol per day.

These prediction variance estimates do not, however, take account of the errors in estimating the regression parameters, a^k, b^k, in Eq 1, or in estimating the regression coefficients associated with the nondietary and dietary covariates in the multivariate prediction models. An adjustment to Eq 6, reflecting the errors in estimating these parameters for Eq 1, adds to this equation the term:

$$Var(\hat{a}^k + \hat{b}^k\bar{z}_i^k) = Var(\hat{a}^k) + 2Cov(\hat{a}^k, \hat{b}^k)\bar{z}_i^k + Var(\hat{b}^k)(\bar{z}_i^k)^2. \quad (7)$$

Table 5 gives estimates of the square roots of $Var(\hat{a}^1)$ and $Var(\hat{b}^1)$. The covariances $Cov(\hat{a}^1,\hat{b}^1)$ for each of the models fit are in Table 4 of Appendix B. In general, however, the addition of Eq 7 has only a small effect on the variance estimate in Eq 6, and this may usually be ignored (except in the case of very extreme values for $\bar{z}_i^k$). For instance, in the above example, the addition of this term changes the estimate of

$$\sqrt{Var(\bar{y}_i^{\infty}|\bar{z}_i^1)}$$

from 5.11 × 10^{-3} to 5.16 × 10^{-3} and hence changes the 95% confidence interval from 0.86–3.51 to 0.86–3.53 μmol per day.